

DOCUMENT RESUME

ED 084 856

EM 011 678

AUTHOR McIntosh, N.
TITLE Evaluation Techniques of Multi-Media Learning Systems.
INSTITUTION Council for Cultural Cooperation, Strasbourg (France).
REPORT NO CCC-TE-73-13
PUB DATE 15 Sep 73
NOTE 30p.

EDRS PRICE MF-\$0.65 HC-\$3.29
DESCRIPTORS Educational Accountability; *Evaluation; *Evaluation Methods; *Instructional Technology; International Organizations; *Multimedia Instruction

ABSTRACT

An attempt is made here to pinpoint some of the problems inherent in evaluations of multi-media instructional systems designed for out-of-school use. Subjects discussed include objectives of multi-media educational systems, some major problems inherent in such evaluations, costs and values of evaluative research, and problems of multi-national research. An example of a possible framework for evaluation is also presented. (RH)

ED 034256



SCOPE OF INTEREST NOTICE

The ERIC Facility has assigned this document for processing to

In our judgement, this document is also of interest to the clearinghouses noted to the right. Indexing should reflect their special points of view.

EVALUATION OF MULTI-MEDIA LEARNING SYSTEMS

COUNCIL FOR CULTURAL CO-OPERATION
STEERING GROUP ON EDUCATIONAL TECHNOLOGY

FILMED FROM BEST AVAILABLE COPY

ED 084856

Strasbourg, 15 September 1973

CCC/TE (73) 13

COUNCIL FOR CULTURAL CO-OPERATION

Steering Group
on Educational Technology

EVALUATION TECHNIQUES OF MULTI-MEDIA
LEARNING SYSTEMS

REPORT

prepared by Mrs N. McIntosh
on evaluative research for out-of-school
multi-media educational systems

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRO-
DUCED EXACTLY AS RECEIVED FROM
THE PERSON OR ORGANIZATION ORIGIN-
ATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRESENT
OFFICIAL NATIONAL INSTITUTE OF
EDUCATION POSITION OR POLICY

CONTENTS

1. Introduction
2. The scope of this paper
3. The objectives of multi-media educational systems
4. Two major problems inherent in evaluating multi-media educational systems
 - 4.1 Is the educational system designed for an open or a closed group?
 - 4.2 Direct or indirect teaching?
5. Other problems involved in evaluating multi-media systems
 - 5.1 Does the course have an entry requirement or not?
 - 5.2 The heterogeneity of the student population
 - 5.3 Possible levels of evaluation
 - 5.4 Evaluating the course as a whole or individual components
6. The cost and value of evaluative research
 - 6.1 Costs involved in setting up the course
 - 6.1.1 The investment in the development of the course
 - 6.1.2 The cost of running the course
 - 6.1.3 The variable cost associated with the amount of personal support built into the course
 - 6.1.4 The variable cost associated with the number of students
 - 6.2 The "value" of evaluative research
 - 6.3 The objectives of the decision-maker
 - 6.4 Formative and summative evaluation
7. Internal or external evaluation?
8. Problems involved in experimentation for the purposes of evaluation
 - 8.1 Experimentation in the social sciences
 - 8.2 "Compulsory" experimentation

- 8.3 Natural experimental situations
- 8.4 Experimentation and sample design
- 8.5 Experimentation in relation to intermediaries
- 9. Problems involved in multi-national research
 - 9.1 Sampling
 - 9.2 Comparability of definitions
 - 9.3 The different structures of the educational systems
 - 9.4 The different availability and the different pattern of the media in different countries
 - 9.5 Measurements of knowledge or achievement
- 10. The characteristics of multi-media courses
- 11. The relationship of the evaluator and the course producer
- 12. One example of a possible framework for evaluation
- 13. Conclusion
- 14. References

CONTEXT

This quotation sets the context for this paper.

"In our approach we will make a distinction between 'evaluation' and 'evaluative research'. The former will be used in a general way as referring to the social process of making judgements of worth While it implies some logical or rational basis for making such judgements, it does not require any systematic procedures for marshalling and presenting objective evidence to support the judgements. Thus we retain the term 'evaluation' in its more common sense usage as referring to the general process of assessment or appraisal of value. 'Evaluative research', on the other hand, will be restricted to the utilisation of scientific research methods and techniques for the purpose of making an evaluation. In this sense 'evaluative' becomes an adjective specifying a type of research. The major emphasis is on the noun 'research' and evaluative research refers to those procedures for collecting and analysing data which increase the probability of 'proving' rather than 'asserting' the worth of some social activity Our task in part will be to evaluate the desirability and feasibility of utilising such social research methods for the purpose of evaluation. For the present, we wish only to explain our distinction between 'evaluation', as referring to the general process of evaluating regardless of the type of 'evaluation study' made, and 'evaluative research', as referring to the utilisation of empirical social research methodology for the purposes of conducting such evaluation studies."

Suchman E.A. Evaluative Research. New York: Russell Sage Foundation. 1967.

EVALUATIVE RESEARCH FOR MULTI-MEDIA EDUCATIONAL SYSTEMS

1. INTRODUCTION

It is beyond the scope of this working paper to provide a comprehensive discussion on the evaluation of educational programmes. What the paper attempts to do, therefore, is pinpoint particular problems which are inherent in the special nature of the courses we seek to study, ie "multi-media" systems designed for "out of school" education, which may be utilised in different countries. The paper also endeavours to look at these problems from the point of view of the researcher attempting to provide information as a basis for evaluation, and does not attempt to cover the whole field of evaluation.

Traditionally, evaluation in educational research has been identified with curriculum evaluation, and the "test and measurement" model of evaluation could be said to have dominated the field of curriculum evaluation for many years. However, Eraut (1972), for example, lists as many as 11 alternative evaluative models, all of which have been at some time to curriculum evaluation. It is quite clear that the evaluation of multi-media educational systems cannot even confine itself to curriculum evaluation, and certainly it cannot confine itself to the "test and measurement" model. For the purposes of this paper, we propose to accept the broader definition given by Astin and Panos (1971) of an educational programme as "any on-going educational activity which is designed to produce specified changes in the behaviour of the individuals who are exposed to it. Thus an educational programme could be a particular method of instruction, a single classroom lesson, a complete course of study, a programmed textbook, the environment of a college, a special remedial programme, an apprenticeship or internship, or an entire school system". Within this wider field, we are only here concerned with multi-media educational systems.

Suchman (1967) takes a broad view of evaluation as referring to the general process of assessment or appraisal of value. Astin and Panos (1971) take a more specific view which is now gaining currency in the area of educational evaluation, and appears particularly applicable to the evaluation of multi-media systems where the educational and financial penalties for bad decisions are greater than for conventional educational programmes.

"... it is assumed that the fundamental purpose of evaluation is to produce information which can be used in educational decision-making. These decisions may be concerned with the continuation, termination or modification of an existing programme, or with the development and possible adoption of some new programme".

2. THE SCOPE OF THIS PAPER

It could be argued that this paper should limit itself to curriculum evaluation, and not extend itself beyond that. Alternatively the paper has been described as concentrating too much on formative evaluation. A third commentator wished it to concentrate on "closed" systems not "open" ones, and so on. These differences make clear the difficulty of the task before us. The definition of "multi-media educational systems" can contain on the one hand a limited course internal to one institution which simply used some hardware, and on the other hand a major educational series designed to go out to the whole country on open network television.

Either the course may be using the mass-media or it may be using some sort of media for mass education or it may simply use the media for education. It may be that we should exclude the last of these from this discussion. On the other hand, it is on such courses that evaluators in this field have traditionally concentrated (Campeau, 1972). Of course it is tempting to stick to the "test and measurement" model, and this, of necessity, would limit the scope of our work. But Parlett (1972) contrasts forcible the "tidy" ideal world of the "test and measurement" type of evaluation which he describes as "the agricultural botany" model with the "untidy" reality more usually found.

"The chief deficiency of a testing-type of evaluation is that the restriction and pre-structuring required for it are so formidable Where are the large-scale innovations, with samples of hundreds of students? The colleges who are willing and able to co-operate in some joint experimental venture? The heads of departments who are prepared to countenance dividing students into experiments and controls?"

The "tidy" and "controllable" world of the experimental psychologist, and the natural scientist before him, find less and less parallel in the untidy world which is likely to make up the environment of a multi-media system. It is interesting to note that it is educational psychologists who have concentrated on and found it possible to concentrate on curriculum evaluation, and sociologists and others who are now more concerned with extending the scope of evaluation. For an educationalist within an institution, the system with which he is concerned may be permitted to end at the boundaries of the institution. Mass multi-media systems in particular are unlikely to be so confined. If it was only a question of the additional cost involved, this would be relatively simple. But it is becoming clear already that multi-media systems have social and even political implications. It could indeed be argued that the evaluator should not be concerned with these and should confine himself to the educational content. All the indications are that this argument is no longer acceptable. The Open University in Great Britain provides an obvious example of a multi-media educational system set up with objectives which were both educational and social. The fact that the system is also cost-effective is an additional bonus. As pressure for expanded educational opportunities increase, so does the necessity to be cost-effective. It is not surprising that countries as diverse as America, Iran, Spain and India are looking hard at the contribution that educational technology and Open University type systems can make to solving their educational problems. The causes of this could be said to be semi-economic and semi-political.

Another danger is worth noting at this stage, one that could be described as semi-political and semi-educational. Multi-media systems, particularly if nationally available have more potentiality of centralised control of content and standards. This may be a good thing, or it may be bad. It could lead to raising of standards, or to uniform mediocrity, or to misuse. All of this is likely to have implications for the evaluator that go well beyond that of curriculum evaluation.

The boundary line between what constitutes a research study and an evaluative study often becomes blurred. Research has an important role to play in the overall set up, design and evaluation of multi-media courses. Some of this research may be evaluative in nature, some may not. It is in our interest, at this stage, not to try to draw the demarcation line too clearly. Several points at which research or evaluation of one type or another can make a contribution can be distinguished.

- a. in discovering whether there is a need for a course
- b. in locating, defining and characterising the target group of the course
- c. in pre-testing the course or its components
- d. in providing short-term remedial feed-back while the course is running
- e. in determining whether or not the course works both in terms of
 - the needs of the student
 - the needs of the course producers and conveyors
- f. in determining whether or not the objectives of the resource-provider have been met.

The research techniques which can be brought to bear on these problems are many and varied. They will come mainly from the areas of social, psychological and educational research. It is not our task in this work either to convey the details of the techniques involved or even to make an exhaustive list of all possible research techniques.

What we hope to do is to illustrate by means of case studies the range of problems the evaluative researcher may face, and the range of techniques which he may have successfully or unsuccessfully brought to bear on these problems.

We wish, in particular, to attempt to systematise knowledge of techniques that have proved valuable.

3. THE NATURE OF THE OBJECTIVES OF MULTI-MEDIA EDUCATIONAL SYSTEMS

It is first necessary to comment, albeit briefly, on the nature of the objectives of such systems. Each system will undoubtedly have multiple objectives. Different objectives will be specified by different persons or organisations within the educational process. There may also be different objectives for different components of the course. Some of these may be interdependent - some may not. Some of these may be in conflict with each other. Some may be short-term and some may be long-term. This complex of objectives is likely to be hierarchical in nature, and different values may attach to different objectives. It is not always possible to define all objectives adequately at the start of a programme, even with the assistance of pre-existing lists (eg Bloom's). More importantly, the dogmatic use of the "objectives" approach assumes that the only likely outcomes of a programme are those anticipated by its statement of objectives (Eraut, 1972). Should any unspecified outcomes be ignored when the programme is evaluated? They may turn out to be more significant, either positively or negatively, than the specified outcomes. And all this pre-supposes that the objectives themselves are worthwhile.

Assuming that we are able to determine adequately the objectives of all the parties involved in the decision to set up a multi-media educational system, we may distinguish sets of people with differing objectives.

- a. society as a whole
- b. the institution/group that makes the course
- c. the designer of the course
- d. the institution that runs the course
- e. the student who is the "user" of the course
- f. the employer of the student, if any
- g. other people who may be affected by the student's participation in the course.

Not all of these categories will be present for any particular course.

Implicit in the establishment of these hierarchies of objectives, is the attempt to measure their attainment. This, in turn, implies agreement on the criteria against which attainment of these objectives should be measured. The major distinction to be made is between criterion-referenced and norm-referenced measurements. Sometimes measurements of both kinds may be desirable and may or may not be possible. For example, we may wish to try to measure the absolute benefit to the student or to society as a whole. Alternatively, we may wish to make relative comparisons between alternative methods of teaching the same material, or between alternative types of institutions providing the same courses. There is frequently no absolute yardstick available, and relative comparisons with fellow students may be all that is possible.

4. TWO MAJOR PROBLEMS INHERENT IN EVALUATING MULTI-MEDIA SYSTEMS

Two fundamental distinctions need to be made arising from the nature of multi-media systems for out-of-school education which affect the selection and use of evaluative techniques in this area. These primary problems are:-

1. the nature of the student population. Is it an "open" or a "closed" group?;
2. is the teaching-learning system a "direct" or "indirect" one? Or to put this another way, is there an intermediary in the learning process?

It is necessary to consider these points in more detail.

4.1 Is the evaluative researcher studying an "open" or a "closed" group?

The terms "open" and "closed" have been chosen by the author of this paper to distinguish between two categories of student group, which are discussed below. Do the producers of the course know exactly at whom they are aiming the course? If so, are they able to locate them? The course may be aimed at an "open" group who may choose to follow the course, but do not have to formally enrol or register in any way, and are therefore not necessarily known. Any course using open network broadcasts, either of radio or television, is likely to have an audience of this kind. Alternatively, the course may be aimed at a "closed" group who are known, or at least can be located, since they have to formally register and perhaps pay a fee for the course.

Irrespective of which of these types of groups the course is designed for, the evaluator needs basic information about the characteristics of the students. For "closed" groups, this and subsequent information is relatively easy to obtain, provided the energy and resources are forthcoming.

For the "open" category this background information, and indeed any other information, is very difficult to obtain, if only for the reason that there is no infallible means of locating the students or even of knowing how many of them exist. The problem is exacerbated for the "open" group by the fact that the student may not have the motivation to co-operate that those in a "closed" group have. By actually joining a "closed" group the student has committed himself either financially, personally or socially to such a course. This distinction between an "open" and "closed" group has implications both for the designer and the evaluator in terms of level and of structure which will be returned to later. Of course there may also be eavesdropper "open" audiences for courses designed for "closed" groups. The desirability and/or difficulty of evaluating these is a separate issue.

4.2 Direct or indirect teaching?

The second fundamental distinction between multi-media courses and more conventional courses which needs to be made, is one that affects both the designer and the evaluator of such courses.

In conventional courses, the course designer, lecturer or writer, is usually in direct contact with the student, ie the user. This direct contact has the strength of adaptability and allows an immediate response to the student. The course can, in effect, be individualised. Design problems are less acute, since mistakes can be remedied individually on the spot. Design is therefore less expensive, and time scales for production are shorter. On the other hand, the course is more expensive to run and less easily transferable to other students and other institutions.

From the point of view of the researcher, however, it is more difficult to evaluate. It is true that "inputs" and "outputs" (Astin and Panos, 1971) may be easier to measure, but "operations" - the means to the achievement of the educational ends - will be more difficult. The course could be viewed as, in effect, a purpose-built product for each individual student. No two products will be the same, since the interaction of the tutor and student will always differ, and attempts to measure the "operations", even if the same measuring devices are used, will inevitably be affected by the differing nature of the "operations".

In a multi-media course, in order to utilise the media to full advantage, one seeks and requires economies of scale. These economies may have to be gained at the expense of flexibility. To this extent, evaluation taken on its own might be made easier. This will be the case when large numbers of students are studying the course "directly".

Frequently, however, it is not as simple as this. The other way to achieve economy of scale is by several institutions adopting the same course, which may or may not originally have been designed for this purpose. The institution in this way acts as a type of agent or "intermediary" in the educational process. Yet another "intermediary" may intervene in the form of a tutor, and more complicated still, the course may be utilised in different ways by different tutors in the same institution. This provides a welcome opportunity to mediate between the, of necessity, undifferentiated educational message and the needs of the individual student in order to individualise the learning system. But although the middle-man or intermediary may introduce welcome flexibility into the educational system, he may alternatively, if he is not adequately familiar with the design and objectives of the course, actually impede the learning process. His intervention, and/or the institution's intervention, will certainly make the process of evaluation more difficult, since their goals may not be the same as those of the course designers, and in fact they may not be interested in evaluating the course at all.

We can therefore attempt to categorise these multi-media courses in terms of the possible degree of difficulty involved in their evaluation.

Difficulty of evaluating
a multi-media course

Type of audience

		Closed	Open
<u>Teaching system</u>	Direct	Easy	Difficult
	Via an "intermediary"	Difficult	Very difficult

5. OTHER PROBLEMS INVOLVED IN EVALUATING MULTI-MEDIA SYSTEMS

5.1 Does the course have an entry requirement, either in terms of ability or attainment, and if so is this known or not known to the students?

This is important for several reasons. If the course is designed to be "open" but does, in fact, have an assumed entry requirement, how can the students themselves discover in advance whether or not they match up to this requirement without a costly commitment either of time, morale or money? Similarly how do the evaluators discover the abilities of the students and whether or not they match up to the entry requirements of the course?

If the course has no entry requirement, then the only measure of success can be an absolute measure of knowledge at the end of the course. Whether this is an adequate measure of success or not will depend on whether the original objective was to bring everyone up to a certain standard irrespective of where they started or if it was simply to produce some (maybe minimal) gain in knowledge.

5.2 The heterogeneity of the student population

The very nature of multi-media courses for out-of-school education is likely to imply a heterogeneous student population. If the course is "open", the heterogeneity may not be known. If the course is "closed", it may be known or partly known, but this is not on its own enough. To recognise a problem is not to solve it. The evaluator will at minimum be able to determine some of the more obtrusive characteristics of the student population and may be able to predict with some accuracy which groups of students should be able to benefit from any particular course pitched at any particular level. If courses are likely to be remade, then formative evaluation may be able to make a contribution to this re-making. If the course is not to be re-made, then prior research is even more necessary to attempt to diagnose which students would prove to be unable to benefit. Preferably even before then, the potential student audience should be researched to help determine the level at which the course should be pitched. The more clearly defined the target groups, the easier all these problems are.

5.3 Possible levels of evaluation

We have already touched on the differing objectives which one course may need to meet for different persons or institutions involved in that course. As an increasing number of stages are involved in the educational process, so the multiplicity of objectives is increased. It is prerequisite of educational programmes that they are evaluated at different levels. It is possible to distinguish, at minimum, four different levels.

- b. evaluation of the effectiveness of the whole course
 - in the opinion of the student
 - in the opinion of the educator providing the course
 - in the opinion of the buyer, user or employer of the newly trained product
 - in the opinion of the outside educationalists
- c. evaluation of individual units or blocks of work on the course, eg one week's work
- d. evaluation of the impact or effectiveness of one individual component of the multi-media learning system, eg the radio or the written materials.

5.4 Evaluating the course as a whole or individual components

For multi-media courses it is, in addition, particularly necessary to attempt to evaluate the contribution of the individual components of the course. Obviously for the whole course to be optimally effective, it is desirable for each individual component of the course to be individually effective. To this end it is possible to invoke a wide variety of research techniques, to pretest broadcast programmes for impact and effectiveness, to pilot written materials for clarity and comprehension, etc. Each component should be pretested individually, Romiszowski (1972) argues that each component of the courses should be tested against its relevant objectives. He suggests that it is preferable to compare media against specific types of learning objectives, and that if one does this, one is automatically forced to evaluate the individual components of the course rather than the course in totality.

It is obviously much easier to accept his view. The problems involved in pretesting the overall effectiveness of the whole educational process are formidable. Even to accept his, less ambitious, view poses problems of a high order, which he discusses in the following terms (Romiszowski, 1972).

This procedure necessarily consumes time and money. In the case of broadcast materials (particularly films and TV) the cost of this exercise is particularly high, and problems peculiar to the medium are encountered.

One problem is to 'simulate the operational environment'. It is generally financially out of the question to prepare, say, a fully complete colour film, try it out, find certain defects, so go back to the studio, replan, rescript, re-engage actors and reshoot certain sequences. Even if finances were available, the time-scale to which most broadcast production teams work would not allow for such a procedure. Any evaluation of training film must take place at an earlier stage of production. Some training film units developmentally test certain important sequences of a programme, by shooting trial runs on videotape, and evaluating them individually before shooting the whole film. Others argue that a low-budget, probably black and white, videotape sequence, rushed off in a hurry, is so unlike the final polished product, as to diminish the value of such evaluation procedures.

Another problem is to locate audiences of suitable 'guinea pigs' for evaluation. Broadcast production teams seldom work within a school or other institution where there is a steady supply of eager students. Also for a proper evaluation, the students must be at the appropriate point in their course. This generally would mean that programmes destined to be used this academic year should have been evaluated and improved during the previous year."

Obviously we do not yet have the perfect formula for evaluating the composite effects of a multi-media course, but we do know enough to consider that to take such an elemental view of evaluation may be dangerous. The model will inevitably be a complex and interrelated one:

- we cannot assume that all media work, or communicate in the same way and to the same extent to different people;
- we cannot assume that all people are equally able to learn from each of the different media;
- we cannot assume that all messages can be conveyed equally well through each medium.

All of this puts on one side the more practical problems of different production time scales and widely varying costs for different media. Millions have been spent by commercial researchers in the field of advertising research in their attempts to systematise their choice between the media. They have not yet found the answers. Clearly the resources available in this field are likely to be more limited.

The main problem is that we do not know whether or not the media work in isolation or in co-operation or even antagonism to each other. Is their effect cumulative, and if so at what sort of rate? Is it an arithmetic or a geometric progression, for example, or does one medium on its own convey say 60% of a message effectively, and the second one add only 20%? The other side of this problem is that individuals may have different learning styles or abilities which cause them to learn differently from each of the media. Some people, for example, tend to be visualisers while others are verbalisers.

How do we evaluate in this complex situation? The evaluation of a method of teaching could be done by manipulating the media, the learning situation, or the message. We could experiment with the same message using different media, but it is impossible to ensure that the same message is conveyed equally on different media. Even if we could ensure this, it would be difficult to define what would be the equivalent impact of the same message on different media. In addition, the level of exposure to different media varies from individual to individual. And short term effects may be different from long term effects. Marc and Durand (1971) suggest that:

"the principal error in comparative studies is to set, a priori, the media as competitors, when they are most often used in a complementary manner: the source of this error is found in the emphasis put on the concept of 'effects' to the detriment of the concept of 'communications'".

In one way, this makes life easier for the course designer, and for the evaluator who no longer needs to seek equivalent standards. On the other hand, it has implications for the course conveyor and the student who may not have all the media equally available to them and may, therefore, miss out on some parts of the message. To evaluate, we need to try to estimate what proportion of each medium has been used by different individuals.

The learning activity as we have said, is likely to be of a different nature with different media, eg print and TV. So if different students have learnt in different ways, can one then use the same measure to estimate the amount of their learning? If, because the learning activities are different, we have to use different measures, how can we ever know if the results are comparable?

6. THE COST AND VALU OF EVALUATIVE RESEARCH

6.1 Costs involved in setting up the course

Costs involved in the actual setting-up of multi-media courses have three main components:

- the investment in the development of the course itself;
- the basic cost of running the course each time it is put on;

- the variable cost associated with the number of students taking the course. The variable cost will depend not only on the number of students studying each time the course is run, but also on the amount of personal contact with academics that is built into the course. At the one extreme it may be nil in an "open" course, or it may include only postage, at the other extreme it may include regular personal contact, and therefore be quite great.

The cost of evaluative research needs to be looked at differently in relation to the three different costs listed.

6.1.1 The investment in the development of the course

Inevitably, different media have widely differing costs. The more expensive the medium the more it is necessary to know how the medium works and how effective it is. Heavy development costs of particular media can only be justified either if the educational gain is overwhelmingly greater than an alternative method, or if the course is likely to reach such a large audience, or be repeated so often that the set up cost is spread over such large numbers over time as to be justified. Exceptions to this statement have been argued on the grounds, for example, of the motivation provided by television. This is not yet proven. A more difficult argument is the quasi-prestige one, that a course looks like a poor relation unless it contains television. If a course is to be run once only, then the development cost will be great in relation to its other costs. Theoretically it would make sense to research into it at this stage. On the other hand, if the course is not to be re-run, then no action is likely to take place as a result of the evaluation and there is therefore no point in carrying it out.

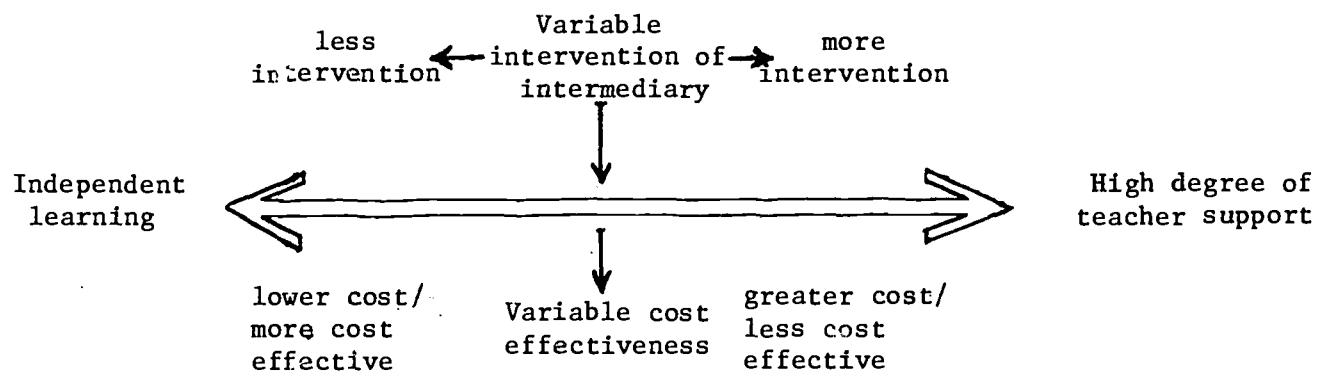
Additionally, heavy set up costs may not be justified for subject matter which is likely to become out-of-date soon.

6.1.2 The cost of running the course

This will be a recurring cost and money spent on researching the operation of the course to reduce this recurring cost may well be cost-effective, particularly if the course is flexible enough to allow amendments to be made before it is re-run.

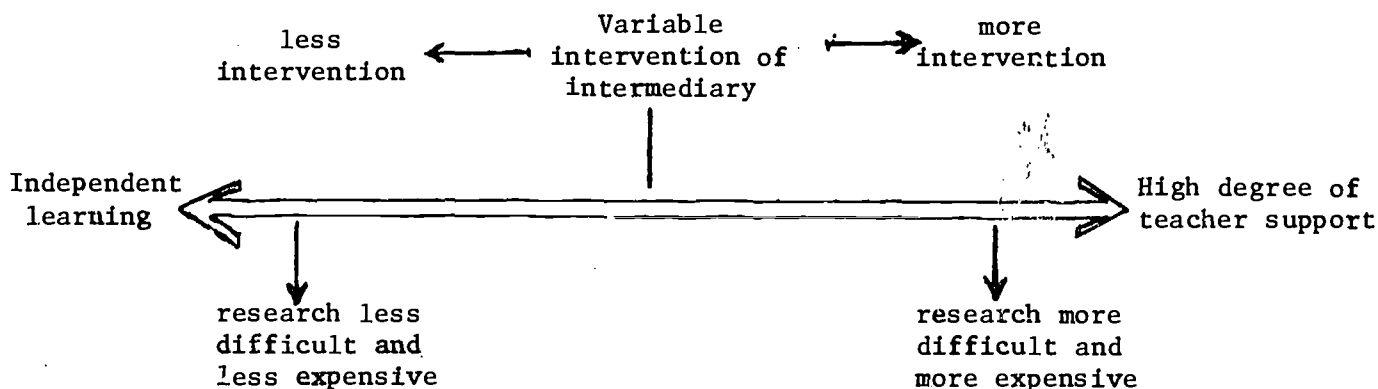
6.1.3 The variable cost associated with the amount of personal support

It is helpful to look at this along a continuum, from the entirely independent learner at one end, to the teacher-supported student at the other. Putting on one side the educational implications, the more an intermediary intervenes, the more expensive the course is likely to be.



The amount that is worth spending on evaluation of this area of cost is likely to depend on the proportion of the cost of the course that is "variable" cost, ie the more personal support that is provided, the more this cost is likely to be. At the same time the research will become more difficult and more expensive.

It is possible to look at the continuum in relation to research in the same way:



6.1.4 The variable cost associated with the number of students

When one course can meet the needs of large numbers of students without too great a degree of intervention, then the economies of scale are likely to be sufficient to outweigh the increased set-up costs. But as evaluative research does its job better, and more is known about the target audience, so the defined group is likely to become smaller, more specific and less replicable. As this happens, the relative advantage of the media, and of multi-media systems, that of communicating economically to large numbers of students in different places and at different times, becomes less marked in comparison to conventional systems. This is particularly likely to be true as the courses get more advanced and more dependent on previous work. It will be ironic but very useful if our increasing efficiency in evaluative research proves that many multi-media courses are not effective either at a particular educational level, or at the available cost level for the size of the audience. It will, however, be in the best tradition of applied research if we can so specify our objectives and analyse the components of our problem that we are able to decide in advance when it is or is not appropriate to devise multi-media courses for specified groups or on specified subjects. Evaluation is conventionally retrospective. Its scope should be extended, since it will ultimately only justify itself if it is able to build up predictors of success or failure and thus avoid investment in poor or not-needed courses.

6.2 The "value" of evaluative research

Since, as we have defined it for the purpose of this paper "evaluative research" is applied research rather than "fundamental" or "basic" research, we are involved in an attempt not to extend the boundaries of knowledge, but more to use our existing knowledge as an aid to the solution of some given set of problems. We

are involved in a systematic search for, and analysis of, information. Green and Tull (1970) consider research as a "cost-incurring activity whose output is information of potential value to management decision".

They state:

"It is apparent that information can never be available to the extent that the decision-maker would desire if no costs were involved. Since obtaining information is a cost-incurring activity, rational decision-making necessarily involves consideration of the value of information. The amount of information is important only as it affects the value The value of this information can be measured in terms of its use in reducing the costs of uncertainty which are associated with taking action based on an earlier (and less) information state."

It is interesting that this view of the value of applied research only developed significantly in the sixties and then mainly in relation to the field of business administration. Hemphill (1969) makes this point, and considers that "evaluation studies in education can be viewed more appropriately within a context of decision-making than within a framework provided by the purposes and conventions of research". He states

"Evaluation studies are made to provide a basis for making decisions about alternatives and, therefore, in undertaking an evaluation study, one at once addresses himself to questions of utility. It may be objected however, that this is too idealistic a view of the purpose of evaluation studies."

What is true is that this view is certainly not the view widely held by the majority of persons currently involved in evaluation studies. Often it must be said, evaluation is seen as an end in itself and not a means to an end. It is not unreasonable to suggest that the view stated by Hemphill, and found already to be valid in industry, may be increasingly relevant to evaluative research on multi-media educational systems where the investment of resources is much greater, where the potential audiences are much larger, and the penalties for bad decisions resulting in bad courses and wasted resources are very high.

6.3 The objectives of the decision-makers

The analogy with business decision-making falls into difficulties, however, at one critical point, that of the decision-maker. McIntosh (1973) discusses the different characteristics of decision-makers in the public and private sectors, and notes that financial objectives may not be the only ones that need to be met. Other objectives are less easy to quantify. Some objectives may in a sense be irrelevant, for example motivations of personal prestige. Particularly for educational systems, "notions of utility" become much more complex and difficult to quantify since social and political concerns may well intrude on educational and economic concerns.

6.4 Formative and summative evaluation

Scriven (1967) distinguishes between formative and summative evaluation. He describes "formative" evaluation as conducted in conjunction with the development of new educational programmes, and "summative" evaluation as that which is used to assess the effectiveness of existing programmes. These two types of evaluation are, of course, complementary to each other, and wherever possible if both are to be carried out should be designed as part of one overall programme. However, if we accept the purpose of evaluation as defined by Astin and Panos (1971), given earlier in this paper, then there may well be occasions when it is inappropriate to carry out summative evaluation.

This may seem heretical, in the sense that one could argue that it is always important to know what has happened. On the other hand, if a course is not going to be re-run, ie if no action is likely to take place as a result of the evaluative study, then clearly the study should not take place. In cases such as this, formative evaluation may be all that is worth doing, using evaluative data collected during the developmental stage of a programme. If, as with television, the initial investment cost is very great, then formative evaluation becomes even more important. To gain the most advantage from formative evaluation, it is better not to develop a large number of small programmes, but to try to develop a fairly small number of large programmes each of which can be properly tested and will be widely used and re-used.

7. INTERNAL OR EXTERNAL EVALUATION?

By internal and external evaluation we mean here whether the evaluation is conducted by the institution who is producing and/or running the course or by some external agency. Since evaluation implies "the social process of making judgements of worth" (Suchman, 1967) it is often argued that it should be carried out by persons who are external to the programme, thereby presumably guaranteeing objectivity. This argument has become increasingly important as large sums of governmental money have been invested in educational programmes, many of which are multi-media in kind. At the same time the task of the evaluator is made more difficult. If he is charged with formative evaluation, it may be difficult for him to be closely enough involved with the organisation to plan and carry out the formative evaluation effectively, to do it in time, rather than too late, and ensure that its results are not just fed back but acted upon. The fact that he is an outsider already creates some problems but he carries the additional burden of being known to be the agent of the funding body, with all that that implies. It could be argued that it is virtually impossible for external evaluators to do formative evaluation well.

Summative evaluation, then, the final judgement of worth, is a different matter. Here it is possibly more important for the evaluator to be objective, particularly if large sums of money are involved. However, without adequate formative evaluation, it is likely that researchers attempting summative evaluation may find themselves without the necessary information on which to work.

How valuable then is external evaluation? To be effective, it must be closely involved with the design and implementation of the system. If it is closely enough involved to be effective, then can it still be objective? One thing that does emerge from this is that the political imposition of external evaluation, for example as a safeguard for taxpayers' money, is likely only to give a spurious air of academic respectability to the project and not to be of any real benefit to the people involved. Under these circumstances it is increasingly likely to become an end in itself and not a means to an end. This is particularly so if the funding is also external. If an institution has committed its own resources, and decided upon evaluation itself, then it is more likely to take notice of the results. Of course an institution can use its own funds to commission external evaluation. That is a different matter. It is the role of the external evaluator externally funded which is the most difficult.

8. PROBLEMS INVOLVED IN EXPERIMENTATION FOR THE PURPOSES OF EVALUATION

8.1 Experimentation in the social sciences

One problem is common to all attempts to experiment in the social sciences and will therefore be touched on only briefly. Any "professional" or scientific evaluation will require controlled conditions that we are rarely able to produce in the social sciences. For example, we would need to know whether the learning we were measuring was the only learning going on at the time, or whether other learning was happening simultaneously. Controlled experimentation would imply a homogeneous group of learners, all of whom we knew could learn in the same way. In fact, of course, we know that people do not learn in the same way. It would be virtually impossible to cause everyone to learn in the same way, even in a one-medium learning situation. With a number of media it would, as we have seen, be similarly impossible to determine what proportion of each of the media had been "used" by different individuals. The alternative would be to artificially prescribe how the multi-media package should be learnt from. This prior prescription would in itself be dangerous, since it might pre-judge the issue of how the package should be best used and so preclude it from succeeding in its objectives. Problems of this nature are inherent in all experimentation in the social sciences, and evaluators would do well not to pretend to a degree of accuracy and control that they are never likely to achieve. These problems are particularly intractable when one is dealing with an "open" student population.

8.2 "Compulsory" experimentation

Another problem is of a different nature, but no less important. Assuming that we accept that some form of experimentation is possible, then this implies alternative forms of educational treatment and provision. With an "open" group who are not necessarily studying for qualifications, some form of course experimentation is possible, providing the first prerequisite of identifying different groups within an "open" audience can be achieved. With a "closed" group, although it is administratively easier to experiment and control, if the students are studying for some form of certificate or a degree, it may be morally unacceptable to offer

different sets of educational experiences to different students, as one package may turn out to be much less effective than the other and that set of students may be materially disadvantaged through no fault of their own. In a situation like the Open University, for example, this solution would be unacceptable as it might affect students' chances of obtaining a degree.

8.3 Natural experimental situations

A further possible alternative would be to seek natural experimental situations, for example, where one part of a population has access to TV and another does not. This is again difficult, as it is highly likely that those people without TV are already different for entirely separate and perhaps relevant reasons, from those who have it, ie they may be a peculiar section of the intelligentsia, or simply people living in remote areas. In this case, these factors would be likely to obscure the other ones that we were attempting to measure.

8.4 Experimentation and sample design

It is worth looking in detail, to illustrate the problem, at just one area that would be involved in any experimental design - that of sampling. We would inevitably require matching samples that would have to be selected in some way, either randomly or purposively.

For the purpose of random sampling, we would need a sampling frame or list of the persons involved from which to select our sample. By definition, with an "open system" such a list is unlikely to be available. If we adopted purposive sampling of some sort, then again with an "open" group this would assume prior knowledge of the population under study that we might not have. Purposive sampling may well be difficult even for "closed" groups since to sample effectively in this way requires prior knowledge of the relevant characteristics of the subjects to be studied. This assumes two things - firstly that we have prior knowledge of the characteristics that we wish to select by, and secondly that we know these characteristics to be relevant. If they are not relevant, then we may increase rather than reduce error. Often one of the main objectives of the project itself is to discover just such relevant variables and to presume them in advance may nullify the whole exercise. Alternatively, we may know which variables are relevant in learning terms, but the researcher may not be able to describe them in operational terms or the interviewer to recognise them, eg motivation or different learning abilities. Most samples are matched of necessity on recognisable objective criteria, such as sex, age, etc which may have little relevance for our purpose.

8.5 Experimentation in relation to intermediaries

As we have said earlier, a multi-media course contains different components which may all be used to convey the educational message. Different courses will use different media, in different combinations. We are here taking a broad definition of media to include not only TV, radio, print and other audio-visual aids, but also personal support by teachers or lecturers.

It is likely that a multi-media course with personal contact will operate differently from a course without any. When an intermediary exists (see paragraph 4.2) then experimentation becomes even more difficult to control since

more uncontrolled variables are introduced. On the one hand the media may come between the teacher and his class. On the other hand, the teacher comes between the media and the student. Which is more important and how, in particular, can we assess the impact of many different individuals interpreting the message to their students in different ways?

It is perhaps possible to evaluate by experiment the conveying of a concept, but not the complex interrelationship of a whole course.

Unless we know whether the media act separately or whether they act to reinforce each other, or maybe even are counter-productive, attempts to set up instruments of evaluation are of little use. Multi-media systems are inevitably complex. To be effective they must also be flexible and therefore the components will interact. Evaluation must embrace this complexity if it is to have any real value.

9. PROBLEMS INVOLVED IN MULTI-NATIONAL RESEARCH

There are particular problems involved in attempting to set up a scheme to classify evaluation techniques that would be valid across different countries.

Some of these are listed below:

9.1 Sampling

- a. When intermediaries are involved, unless we know in advance which institutions are running course, it may be difficult to obtain comparative samples of institutions in different countries, since their educational systems may not admit of easy comparison.
- b. When we wish to sample cross-sections of the population direct, the normal problems of comparability apply. Are lists of the same nature available in all countries? Are they adequate for the purpose? Are they lists of individuals or households, etc? For example, area sampling is widely employed for individuals and households in America, where as in England such samples are usually obtained from the Register of Electors.

9.2 Comparability of definitions

Almost every country defines almost every concept with which we are concerned differently for its own purposes. To take one simple example - social class. The apparent proportion of different social class groups who go on to higher education varies widely across countries. When this is studied more closely it is apparent that the definition of social class also varies widely.

9.3 The different structure of the educational systems

This follows on immediately from the previous point. Educational systems vary widely between countries both with respect to age of entry and exit. Different institutions exist at different levels for different purposes. Comparisons of entry behaviour based on national criteria, for example, may be impossible.

9.4 The differential availability and the different pattern of the media in different countries

Some countries have a tradition of state-owned television and radio, some have commercial television as well. It is possible that people's reactions to learning from radio and television in particular may be affected by the different degrees of commercialisation, for example, in their existing broadcast output.

9.5 Measurement of knowledge or achievement

The problem of developing diagnostic tests or tests of achievement that are valid for people of different linguistic, cultural and educational backgrounds is well known.

This is obviously not a comprehensive list, and the issues raised in 9.1 - 9.5 all interact but it suffices to give an idea of the scale of the problems involved.

10. THE CHARACTERISTICS OF MULTI-MEDIA COURSES

It is necessary, therefore, to characterise multi-media courses in terms of their key features before proceeding further in an attempt to classify appropriate evaluative research techniques.

The following set of features, though not exhaustive and not mutually exclusive, provide a basis for further discussion.

1. Is the course aimed at an "open" or "closed" student population or both?
2. Is the learning system "direct" or via an intermediary?
3. Is the course to be repeated, or is it to be run once only?
4. What is the relationship of the cost of the course to the cost of possible evaluation procedures?
5. Does the course have an "assumed entry behaviour" or not? If yes, then is it known, and if so, by whom?
6. If the course has an "assumed entry behaviour" is it desirable to measure it or not?
7. Is the performance of the student to be assessed or not?
8. What combinations of media are to be used, and in what ways?
9. Is the course to be evaluated unit by unit; (week by week), or as a whole or both?

Taking the first two sets of features as basic provides a matrix which enables us to characterise the four main types of courses which are likely to confront the evaluator.

Type of audience

<u>Teaching system</u>	<u>Type of audience</u>	
	Closed	Open
Direct	1. Closed and direct	3. Open and direct
Via an "intermediary" (mediated)	2. Closed and mediated	4. Open and mediated

These categories of courses are not mutually exclusive, and some courses may be designed for both types of audience and/or for both types of teaching system. Yet again, some courses may be designed for one type of audience, but also be used, by chance or as a bonus, by the other type. The "eavesdropper" audience of academics and other interested persons who watch Open University programmes but are not registered students is an example of this.

It is sufficient here to give an idea of the sorts of courses which are likely to fall into these four main categories.

1. Closed and direct - programmed learning, eg using teaching machines.
- educational TV or radio for classes with no "professional" teacher involved.
2. Closed and mediated - Educational TV or radio in colleges, mediated by a "professional"
- The Open University
3. Open and direct - RTS promotion
- Teleac
- BBC Further Education for Adults
4. Open and mediated - The joint course of the Finnish Broadcasting Corporation and the National Board of Schools for Inservice Training of Teachers.

Another element which will vary from course to course and will need to be considered by the evaluator is whether or not the course has been designed to include some form of two-way interaction between teacher and taught. In addition to this the course may include the possibility of one-way feedback. This feedback could as well be designed for informational ends as for educational ends, although interaction is more likely to be related to educational ends.

It is quite clear that the majority of studies considered by Campeau (1972) fall into categories one and two. Her review emphasises the "appalling lack of recent objective data on the instructional effectiveness of audio-visual media particularly in the field of adult education". The review excluded by its own definitions and the criteria it adopted much of the literature on the use of instructional media to teach adults that was published in that period.

Campeau (1972) states:

"Most of this non-experimental literature consisted of surveys, testimonials, historical and descriptive assessments, reports of informal evaluations - all of which did not even attempt to deal with or assess the instructional effectiveness of audio-visual media. Instead, assessments were made of user preference for and attitudes toward various media, patterns of media use, characteristics of the post-school audience, problems of educating adults via mass-media and so on. Much of the data offered was in the form of questionnaire responses, enrolment and completion statistics, cost figures and tallies of services provided and extent of use."

The difficulties of experimentation for the purposes of evaluation have already been discussed (8.1 - 8.5). It is quite clear that the only categories in our matrix that are amenable to the imposition of experimental conditions are the two on the left hand side (1 and 2). It is no coincidence therefore that Campeau's screening criterion "studies in which adult learners were assigned to experimental and control groups that included, normally, at least 25 subjects each" produced in the main, studies fitting into these categories a form of self-fulfilling prophecy.

What needs to be considered is whether or not it is constructive to so limit the field. It could be argued, as I hope this paper has started to do, that the problems involved in the evaluation of multi-media courses are too complex to be amenable to test and measurement devices under controlled experimental conditions. Some part of their content may and probably must be so measurable, but to assume that all of it is may be to mistake the wood for the trees.

This is not to put on one side the contribution of the psychologists to testing the actual effectiveness of the educational message. It is to say, however, that for multi-media educational systems, other problems may be so important as to change the whole perspective of the evaluator's activities. One example suffices - if an adult worker is not at home at the time of the broadcast, or from experience has decided that TV is a non-assessable part of a course, then the fact that television, after formative evaluation, etc, conveys a concept "perfectly" is totally irrelevant since the student for totally different reasons will not have been exposed to the educational message.

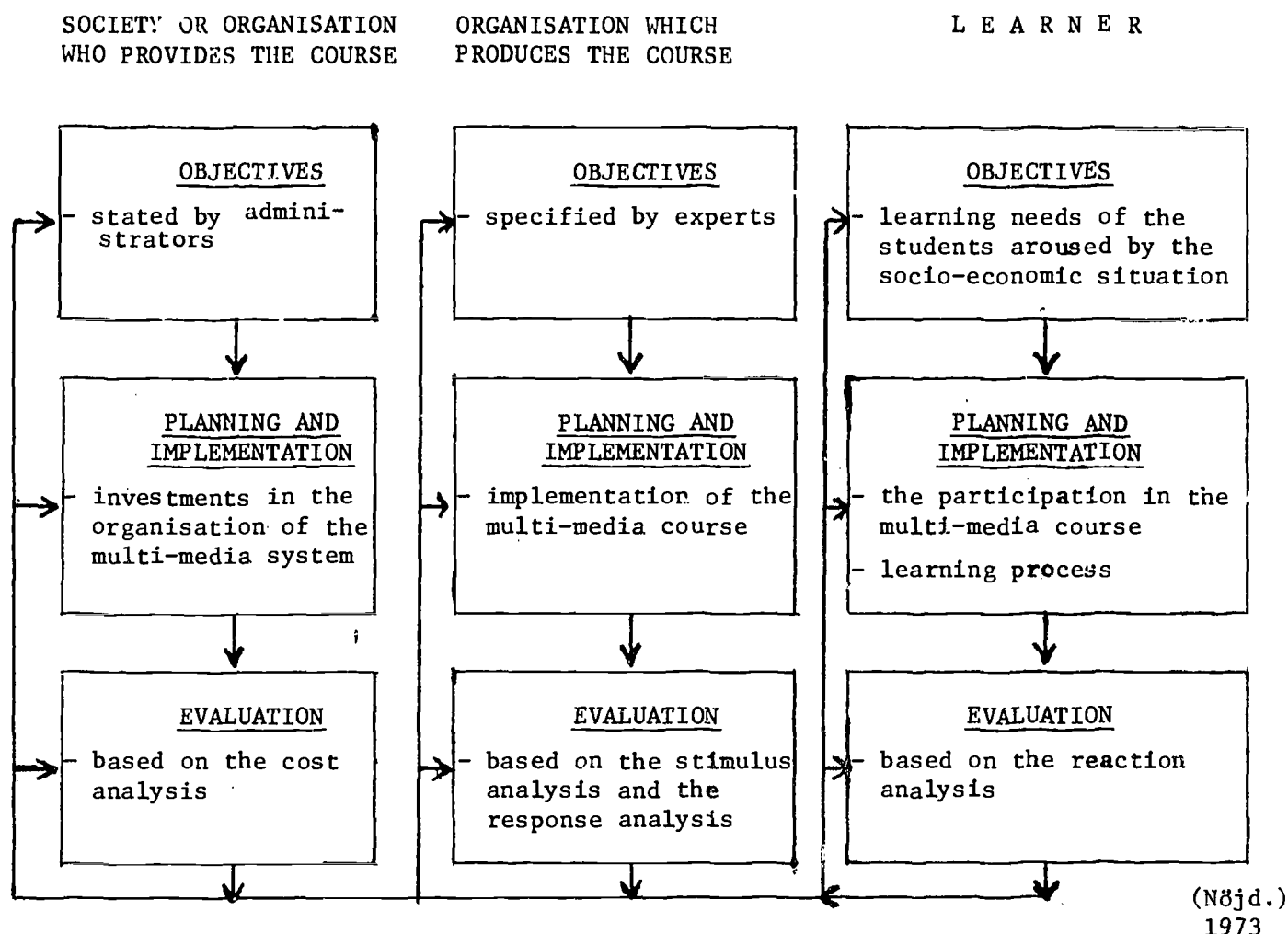
11. THE RELATIONSHIP OF THE COURSE PRODUCER AND THE EVALUATOR

If evaluative research is to be carried out adequately there must be close collaboration between researcher and course designer at all stages. Glikman (1972) makes this point strongly.

"Au stade actuel de développement des systèmes multi-media d'enseignement, la collaboration dès le début de l'action entre le responsable de cette action et le chercheur est indispensable, afin que le premier puisse envisager, dans son programme, les moyens d'évaluation, et que le second puisse, dans son évaluation, tenir compte des objectifs visés."

Another contributor (Najd, 1973) sees the whole process of evaluation as a feedback loop.

Evaluation of the multi-media course at the different levels



All the possible levels of objectives (see para 3) are influenced by the evaluation process. The feedback information has to be channelled in such a way that the different levels - society, the producers of the course, and the learner are able to interpret and use the data.

At the learner's level, the feedback may affect

- the learning process
- the decision to participate in multi-media courses in the future.

At the level of the institution, the feedback may produce crucial methodological advances.

At the level of society as a whole, decisions can be made on a cost-effectiveness basis about the investment in the development of the course.

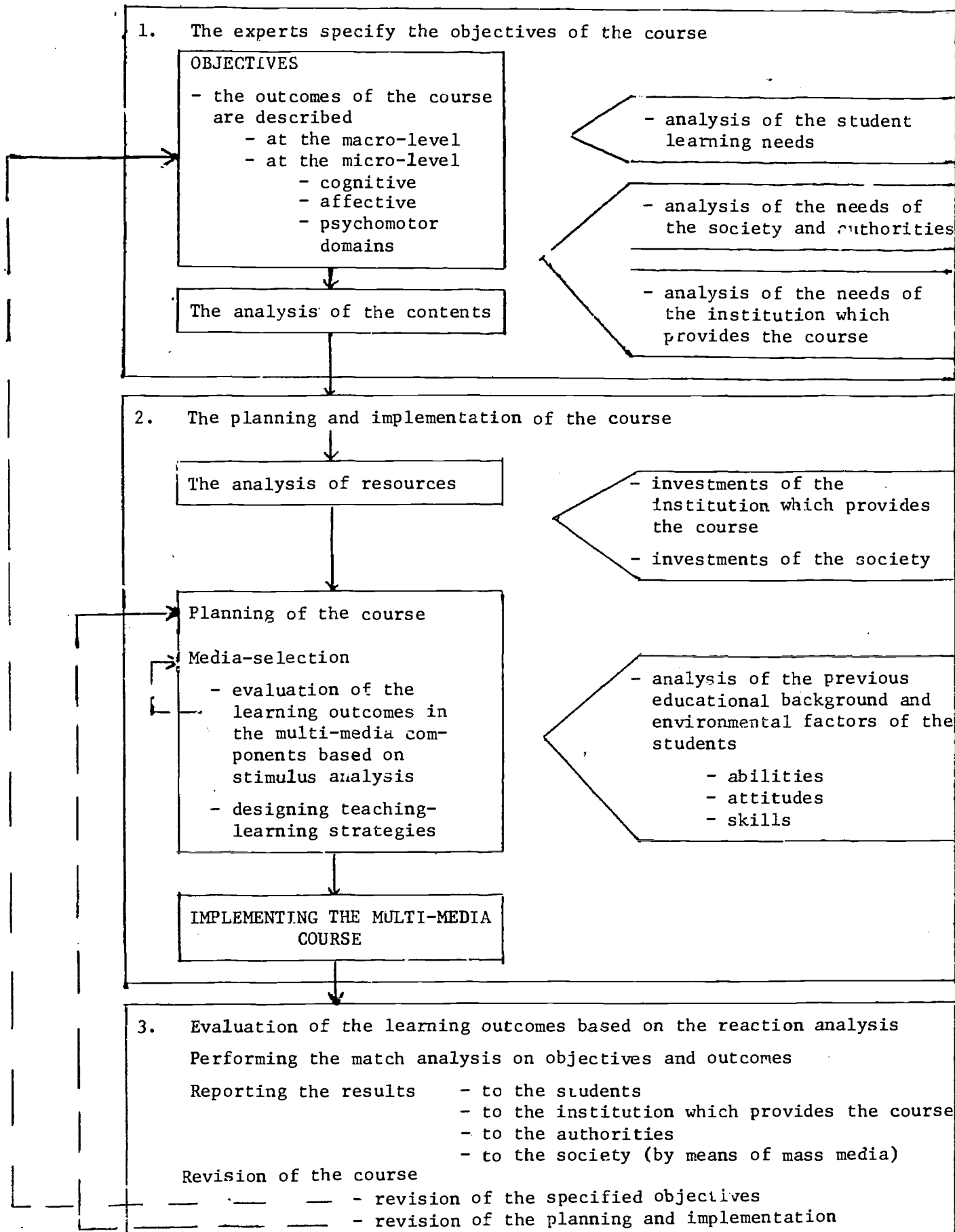
Nojd also outlines the feedback loop in more detail as it relates to the organisation who produces the course. This system appears in Figure 2 (page 34).

12. ONE EXAMPLE OF A POSSIBLE FRAMEWORK FOR EVALUATION

One classification prepared by Glikman (1972) which attempts to embrace all these complexities is included here as a basis for further discussion. (Figure 3, p. 25). This classification, based particularly on the case of open network broadcasting, distinguishes a priori evaluation - before the broadcast, and a posteriori evaluation - during and after the broadcast. This assumes a course in the right hand top corner of our matrix with no mediation from an intermediary, when the broadcast is the key message.

Glikman's framework is quite clearly designed to go beyond the narrow experimental design of much traditional evaluation. Although Campeau's review looked at research among adults, since its criteria were narrowly defined, the majority of the research it cites is among adults learning in relatively conventional situations, for example - college students. It is highly likely that the learning situation for out-of-school education will be less conventional particularly if multi-media systems are utilised. If it is less conventional it is likely to be less controllable, and probably more varied.

Evaluation of the multi-media course at the level of the organisation which produces the course



13. CONCLUSION

The different approaches taken by Glikman and Nojd give valuable insights into the problem of trying to systematise the use of research techniques in such a wide field. It is possible that Nojd's outline stems more from a psychological approach developing from the more established educational tradition. Glikman's on the other hand may stem more from a sociological approach, developing from the problems of broadcasting.

Maybe the field is too wide and we are endeavouring to set up some systematic framework for things that are not intrinsically similar. It could be argued that there is no necessary connection between organisations such as the BBC, Teleac and RTS which use the mass media for educational courses for adults, and individual colleges or institutions who happen to choose some sort of audio-visual aids instead of conventional learning systems for their courses. Any attempt to find similarities may be counter-productive as the dissimilarities are too great.

What they have in common, and this may not be enough to justify the whole exercise, is that they are trying to construct a situation in which one educational message can be used economically and effectively for large numbers of people, or for people who would otherwise not be able to receive it. And this one message has to meet the differing needs of all these people.

Under these circumstances no one prescription for evaluation is likely to meet the needs of all kinds of courses. It may be helpful then to look at the four categories delineated in para 10, and see what lessons can be learnt from previous attempts to evaluate them. We propose, therefore, to look at these categories in greater detail. In order to do this we shall select examples of courses which come into each of these categories, make a study of the problems involved in evaluating them, relate these to the framework on which we have decided to work, and analyse the research techniques which have been found valuable or have not succeeded.

Our objective is not to provide a complete dossier, with instructions of all available research techniques. These are well-documented in existing texts. We aim by adopting a case-study approach to choose cases which not only illustrate different types of evaluation but also exemplify different research techniques.

la diffusion ou "a priori"	Programmation (1)	<p><u>Evaluation de la nécessité d'un cours</u> Détermination de la (ou des) matière(s) à enseigner ("Jusqu'en formation")</p> <p>Q — Besoins des étudiants — besoins exprimés — sondages — besoins mesurés — analyse des programmes scolaires</p> <p>Besoins en main d'œuvre — analyse économique</p>
	Préparation du cours (2)	<p><u>Evaluation de la forme et du contenu à donner au cours</u> (le cours dans une certaine dimension)</p> <p>⇒ Détermination du niveau exact, de la méthode, du matériel, de la manière dont les informations seront présentées</p> <p>Q — Réunions de spécialistes Q — Analyse des programmes scolaires et des connaissances requises par les tests, examens, diplômes,...</p> <p>Q — Sondages et interviews (par exemple, analyse des représentations en rapport avec la matière dans le public-cible)</p>
	Validation du cours (3)	<p><u>Evaluation de la validité des produits</u></p> <p>Q — Pré-testings de produits expérimentaux — émissions-maquettes — documents écrits</p> <p>par groupes — de spécialistes — de formateurs — du public-cible (échantillons représentatifs)</p>
Pendant et après la diffusion ou "a posteriori"	Evaluation de la structure du public (4)	<p><u>Evaluation du nombre et des caractéristiques socio-démographiques du public et de l'évolution</u> (sexe, âge, catégorie socio-professionnelle, lieu d'habitat, situation de famille, niveau d'études, études en cours, ...)</p> <p>⇒ Utilisateurs de plusieurs media — isolés ou regroupés —</p> <p>Q — questionnaires — au moment de l'inscription ou de la demande des documents écrits — joints aux documents écrits — envoyés par correspondance</p> <p>⇒ Téléspectateurs seulement</p> <p>Q — sondages sur l'ensemble de la population touchée par le réseau de diffusion</p>
	Evaluation des motivations et attentes	<p><u>Au début de la diffusion d'une série</u></p> <p>Q — Interviews auprès d'échantillons de public décidé à suivre (demandeurs de documents d'accompagnement — inscrits)</p>
	Evaluation de la satisfaction déclarée (5)	<p>retours chez les mêmes enquêtés <u>en cours ou en fin de diffusion</u></p> <p>Q — Interviews — opinions — assiduité ou abandons — analyse des causes d'abandon</p>
	Evaluation des attitudes et du mode d'utilisation (6)	<p><u>Opinions exprimées — Comportements d'écoute et de travail sur les documents écrits</u></p> <p>⇒ Isolés utilisateurs de plusieurs media</p> <p>Q — Interviews à domicile Q — Réunions de groupes Q — Panels Q — Questionnaires — sur l'ensemble de la diffusion — séries ou émissions — documents écrits — sur des produits particuliers</p> <p>⇒ Regroupés</p> <p>Q — Questionnaires Q — Observations de séances de travail Q — Entretiens avec responsables des groupes</p>
	Evaluation de l'acquisition de connaissances (7)	<p><u>Mesure de l'apprentissage</u></p> <p>Q — Interviews détaillées — mais surtout</p> <p>Q — Exercices programmés ou non corrigés sur ordinateur Q — Exercices corrigés traditionnels Q — Examen</p>
	Evaluation du "renforcement" (8)	<p><u>Fonction de l'objectif de départ</u></p> <p><u>Niveau individuel</u></p> <p>→ modification des attitudes et des comportements (comportements éducatifs, par ex.) si objectif = sensibilisation, incitation, ...</p> <p>→ capacités accrues de réalisation si objectif = apprentissage</p> <p>Q — Interviews Q — Suivre individus ou groupes touchés par les enseignements</p> <p>⇒ Modification des comportements quotidiens — éducatifs — professionnels — sociaux, ...</p> <p><u>Niveau global</u></p> <p>Q — Analyse du système et de ses fonctions dans le cadre — de l'ensemble du système éducatif — de l'ensemble du système global socio-économico-culturel et de son évolution</p> <p>Q — Mise en rapport des résultats avec les objectifs et hypothèses de départ</p>
	Evaluation comparative des différents media (9)	<p><u>Analyses aux niveaux (4), (5), (6), (7) et (8) en fonction des media utilisés par les individus touchés par le système</u></p>
	Evaluation comparative de l'efficacité du système multi-media par rapport aux enseignements traditionnels (10)	
	Evaluation financière (11)	<p>Etudes économiques sur le "coût de la formation" tenant compte</p> <p>— du nombre d'individus touchés</p> <p>mais aussi</p> <p>— des résultats obtenus à tous les autres niveaux précités</p>

14. REFERENCES

Astin, A.W., and Panos, R.J. (1971)

The evaluation of educational programs, in R.L. Thorndike (Ed)
Educational measurement Washington: American Council of Education,
1971 p. 773 - 751

Campeau, P.L. (1972)

Selective review of the result of research on the use of audio-visual
media to teach adults. Council of Cultural Co-operation, Council of
Europe, Strasbourg, 1972

Eraut, M. (1972)

Strategies for the evaluation of curriculum materials, in K. Austwick and
N.D.C. Harris (Ed), Aspects of Educational Technology. London: Pitman 1972

Green, F. and Tull, R. (1970)

Research for marketing decisions, 2nd edition, Englewood Cliffs,
New Jersey: Prentice Hall 1970

Hemphill, J.K. (1969)

The relationships between research and evaluation studies. In Tyler, R.W. (Ed)
Education evaluation; new roles, new means Chicago: NSSE, 1969

Marc, M. and Durand, J. (1971)

Le choix entre les medias. Paper presented at ESOMAR/WAPOR Congress
1971, p. 349 - 370

McIntosh, N.E. (1973)

The use of survey research in the planning of higher education. ESOMAR
Congress, Budapest 1973.

Parlett, M. (1972)

Evaluating innovations in teaching. In Butcher, H.J. and Rudd, E.
Contemporary problems in higher education London: McGraw Hill 1972

Romiszowski, A.J. (1972)

Notes on the evaluation of multi-media teaching systems. Council of
Europe, 1972

Suchman, E.A. (1967)

Evaluative research New York: Russell-Sage Foundation 1967